Postel (Chrome)

Network Working Group
Request for Comments: 936

Michael J. Karels
UC Berkeley
February 1985

Another Internet Subnet Addressing Scheme

Status of this Memo

This RFC suggests a proposed protocol for the ARPA-Internet community, and requests discussion and suggestions for improvements. Distribution of this memo is unlimited.

Introduction

There have been several proposals for schemes to allow the use of a single Internet network number to refer to a collection of physical networks under common administration which are reachable from the rest of the Internet by a common route. Such schemes allow a simplified view of an otherwise complicated topology from hosts and gateways outside of this collection. They allow the complexity of the number and type of these networks, and routing to them, to be localized. Additions and changes in configuration thus cause no detectable change, and no interruption of service, due to slow propagation of routing and other information outside of the local environment. These schemes also simplify the administration of the network, as changes do not require allocation of new network numbers for each new cable installed. The motivation for explicit or implicit subnets, several of the alternatives, and descriptions of existing implementations of this type have been described in detail [1,2]. This proposal discusses an alternative scheme, one that has been in use at the University of California, Berkeley since April 1984.

Subnet Addressing at Berkeley

As in the proposal by Jeff Mogul in RFC-917, the Berkeley subnet addressing utilizes encoding of the host part of the Internet address. Hosts and gateways on the local network are able to determine the subnet number from each local address, and then route local packets based on the subnet number. Logically, the collection of subnets appears to external sites to be a single, homogenous network. Internally, however, each subnet is distinguished from the others and from other networks, and internal routing decisions are based on the subnet rather than the network number.

The encoding of subnet addresses is similar to that proposed in RFC-917. In decomposing an Internet address into the network and host parts, the algorithm is modified if the network is "local", that is, if the network is a directly-connected network under local administrative control. (Networks are marked as local or non-local

at the time each network interface's address is set at boot time.) For local addresses, the host part is examined for a subnet number. Local addresses may be on the main network, or they may be on a subnet. The high-order bit of the host number is used to distinguish between subnets and the main net. If the high-order bit of the host field is set, then the remainder of the high-order byte of the host part is taken to be the subnet number. If the high-order bit is clear, then the address is interpreted in the normal fashion. For Class A networks, using 8-bit subnet fields, this allows a network with up to 127 subnets, each of 65535 hosts maximum, and a main net with 2↑23 hosts. Class B nets may include 127 subnets, each of up to 255 hosts, and 32767 hosts on the main net. Class C networks are not currently included in this scheme. They might be reasonably be added, using four bits of the host part for a subnet desgination and four bits for the host, allowing 8 subnets of 15 hosts and 126 hosts on the main net.

The current implementation does not use subnet numbers separately from the network field, but instead treats the subnet field as an extension of the network field. Functions that previously returned the network number from an address now return a network or network-subnetwork number. Conveniently, Class A subnets are distinguishable from Class B networks, although each is a 16-bit quantity, and Class B subnets are disjoint with Class C network numbers. The net result is that subnets appear to be separate, independent networks with their own routing entries within the network, but outside of the network, they are invisible. There is no current facility at Berkeley for broadcasting on the logical network; broadcasting may be done on each subnet that uses harware capable of broadcast.

Discussion

There have been several earlier proposals for methods of allowing several physical networks to share an Internet network designation, and to provide routing within this logical network. RFC-917 proposes a means for encoding the host part of each local address such that the hosts, or the gateways connecting them, are able to determine the physical network for the host. The current proposal is most similar to that scheme; the differences are discussed in detail below.

Another proposal (RFC-925) involves the use of intelligent gateways to perform routing for unmodified hosts, using an Address Resolution Protocol (ARP) [2]. This has the advantage of placing all modifications in the gateways, but is likely to require additional routing protocols and caching mechanisms in the gateways in order to avoid excessive broadcasts for address resolution. A modification of

this method is to perform encoding of subnets within host addresses by convention to simplify the routing in the gateways, without modifying host software to recognize these subnet addresses. These techniques were not considered for use at Berkeley, because all packet forwarding was being done by multi- homed hosts, all of which ran the same software as the singly-homed hosts (4.2BSD Unix).

The most recent proposal, RFC-932 [3], provides subnetting by encoding the network part of the Internet address rather than the host part. Ordinary hosts need not know of this convention, eliminating the need for modification to host software. Gateways would be able to take advantage of this encoding to compress the routing information for the collection of networks into a single entry. Unfortunately, implementation of that scheme would require a fairly concerted transition by the gateways of the Internet, or the transition period would be likely to overflow the routing tables in the existing gateways. All of the hosts on the larger networks would be forced to change addresses from their current Class A or B addresses to "B 1/2" addresses. There are a limited number (4096) of blocks of Class C addresses available using this encoding. The number of universities and other organizations having already implemented subnets or contemplating their installation argues for a more extensible scheme, as well as one that can be implemented more quickly.

The current proposal is most similar to that of RFC-917; indeed, the two implementations are nearly compatible. There are two differences of significance. First, the use of a bit to distinguish subnetted addresses from non-subnetted addresses allows both smaller subnets and a larger (physical or logical) main network. Half of the host addresses within a Class A or B network are reserved for use in subnets, the other half are available for the primary net. This may useful when using a hardware medium that is capable of supporting large numbers of hosts or for transparent subnetting (e.g. using ARP-based bridges). The corresponding disadvantage is that fewer subnets may be supported. The allocation of bits between the subnet number and the host field could be adjusted, but for Class B networks, neither is excessively large. Given the limited address space of the current Internet addressing, this is a difficult choice.

The second difference is that the width of the subnet field is fixed in advance. This simplifies the already-too-complicated code to interpret Internet addresses, and avoids the bootstrap problem. If the subnet field width is to be determined dynamically, some fraction of the hosts on a network must be prepared to specify this value, and the situation will be unworkable if one of these hosts does not make the correct choice or none are accessible when other machines come

up. Also, the recovery procedure proposed by RFC-917 seems unnecessarily complicated and liable to fail. Dynamic discovery of this value depends on another modification as well, the addition of a new ICMP request. The alternatives are to specify the field size as a standard, or to require each implementation to be configurable in advance (e.g with a system compilation option or the use of a system patch installed when a host is initially installed. The use of a standard field width seems preferable, and an 8-bit field allows the most efficient implementations on most architectures. For Class C nets, a 4-bit field seems the only choice for a standard division.

References

[1] J. Mogul, "Internet Subnets", RFC-917, Stanford University, October 1984

[2] J. Postel, "Multi-LAN Address Resolution", RFC-925, USC-ISI, October 1984

[3] D. Clark, "A Subnet Addressing Scheme", RFC-932, MIT-LCS, January 1985